心理学报 2023, Vol. 55, No. 1, 94−105
Acta Psychologica Sinica https://doi.org/10.3724/SP.J.1041.2023.00094

老龄化对 F0 轮廓线索在汉语言语识别去掩蔽作用的影响*

吴梅红

(厦门大学信息学院; 厦门大学心理学研究所, 厦门 361005)

摘　要　动态基频(F0)轮廓有助于嘈杂环境下的言语识别, 可以作为将目标语音从背景声中分离的知觉线索。本研究通过评估老年人与年轻人在言语掩蔽下聆听具有自然动态 F0 轮廓与对 F0 轮廓操作调节后的汉语语句的言语识别能力, 探讨老龄化对 F0 轮廓线索在汉语言语识别去掩蔽作用中的影响。结果显示, 在言语掩蔽下自然动态的 F0 轮廓比压扁或拉伸的 F0 轮廓更能帮助年轻人抵抗信息掩蔽识别目标言语; 而老年人在言语掩蔽下却难以从动态 F0 轮廓线索中受益。研究结果揭示了老年人利用 F0 轮廓线索促进掩蔽下言语感知能力的老化特点。

关键词　汉语言语识别, F0 轮廓线索, 老龄化, 言语掩蔽, 去掩蔽

分类号　B842; B844

1　引言

噪声下的言语识别及理解对日常生活和工作中人与人之间的交流非常重要。人类具有从复杂的声学环境中提取目标信息并对其进行深度加工的能力。在嘈杂环境中, 人们对特定目标语音的加工会受到周围背景噪声和其他说话人语音的干扰, 这种现象称为"听觉掩蔽" (Auditory Masking)。虽然来自不同声源的声音会在听者的外耳叠加混合, 但听者依然能够选择性地对需要注意的目标说话人所说的内容进行追踪和识别(Schneider et al., 2007)。在充斥着掩蔽声音的嘈杂环境中, 听者会利用各种感知线索将目标言语从背景干扰中分离出来, 从而达到对目标言语的去掩蔽(Unmasking)作用(综述见: 李量 等, 2017)。

嘈杂环境下听者会利用某些与目标言语特征有关联的知觉线索来促进对被掩蔽的目标言语的识别(李量 等, 2017)。基频(F0)轮廓对嘈杂环境下的言语感知非常重要, 它的功能不仅是强调特定的单词和标记短语边界(Cutler et al., 1997), 还可以作为从掩蔽声音中分离语音的感知线索(Binns & Culling, 2007)。近年来许多研究者研究了 F0 轮廓线索对嘈杂环境下言语感知的影响(Binns & Culling, 2007; Chen et al., 2018; Miller et al., 2010; Patel et al., 2010; Shen & Souza, 2019; Wu, 2019), 一致认为 F0 轮廓动态变化对噪声下的言语可懂度具有重要作用。

基频(F0)轮廓变化有助于多语流中的说话人分离, 对日常交流中的言语识别非常重要(Greenberg et al., 2004)。作为不利声学环境下促进言语识别的强有力线索之一, F0 轮廓动态变化能够增强语音流的感知连续性(Dannenbring, 1976), 并提高对背景掩蔽下的目标言语识别(Shen & Souza, 2019)。因此, 许多研究都将兴趣集中在 F0 轮廓线索对言语感知的贡献上(Binns & Culling, 2007; Miller et al., 2010; Shen & Souza, 2019)。研究者通过对自然语句 F0 轮廓的调节操作, 例如压扁(Binns & Culling, 2007)、拉伸(Miller et al., 2010)或反转(如: Hillenbrand, 2003)基频 F0 轮廓, 来检验对不同 F0 轮廓调节后对言语可懂

收稿日期: 2021-08-23

* 国家自然科学基金项目(31771252、32071057、31200769)、福建省自然科学基金项目(2019J01002)和浙江省 CCBD 开放课题(PD11001005002009)资助。

通信作者: 吴梅红, E-mail: wmh@xmu.edu.cn

度的影响。对于英语语音的感知，以往的研究结果显示，无论是在安静状态下还是在稳定的语谱噪声中，通过扁平化英语句子的 F0 轮廓，语音可懂度会降低(如：Binns & Culling, 2007; Miller et al., 2010; Shen & Souza, 2019)。在对这些发现提出的解释中认为 F0 轮廓的上升和下降会将听者的注意力引向言语的内容词；因此，如果没有这些线索，言语的可理解性就会下降(Binns & Culling, 2007)。Miller 等人(2010)也证明了被调节的 F0 轮廓对噪声下的语音识别有害，拉伸的 F0 轮廓与保留自然动态 F0 轮廓相比并没有发现更有利的效果。Hillenbrand (2003)也发现自然动态 F0 轮廓线索对于不利条件下的语音识别更有帮助，语音中的平坦或倒转 F0 轮廓会导致言语可懂度的降低。这些研究结果一致表明自然动态变化的 F0 轮廓有利于英语语音的识别。

基频 F0 轮廓线索对噪声环境下的英语语音感知作用已经得到了广泛的研究(Binns & Culling, 2007; Miller et al., 2010; Shen & Souza, 2019)。对于汉语语音感知的研究，现有的数据集还很有限，但动态 F0 线索有助于听者在背景噪声下言语识别的这种效应在汉语语言中也得到了证实。Patel 等人(2010)使用有意义的汉语句子为目标刺激，证明了在安静的环境中，具有平坦 F0 轮廓的汉语句子与具有自然 F0 模式的句子一样容易理解(Patel et al., 2010)；然而，正如其他几项研究(如：Wang et al., 2013; Wu, 2019)报告的那样，在稳态噪声或嘈杂语掩蔽下，F0 轮廓线索被调节的句子的言语可懂度较低。

在汉语普通话中，F0 轮廓是单音节水平上声调感知的主要线索(Xu & Zhou, 2011)，另外它还与句子语调和音高的感知有关。Wu (2019)使用中文无意义句子作为言语刺激，对句子的语调进行扁平化或拉伸的调节操作，同时保留每个单音节的声调信息。与英语语音感知的研究(Binns & Culling, 2007; Miller et al., 2010)结果一致，具有自然 F0 轮廓线索的语音比对 F0 轮廓线索调节操作后的语音更容易被感知。此外，Wu (2019)的研究还表明压扁或拉伸 F0 轮廓线索会对噪声环境下的言语可懂度产生不利影响，尤其是在多人说话的言语掩蔽下。这些结论说明了动态 F0 轮廓线索有助于言语掩蔽下的汉语言语识别。F0 轮廓线索在汉语言语识别去掩蔽中具有重要作用。

值得注意的是，上述研究都是针对具有典型听力的年轻被试进行的，动态 F0 轮廓线索有助于年轻听者在言语掩蔽下有效的语音识别已经在多个数据集上得到了一致验证，然而关于老年听者的研究还很欠缺。探讨老年人能够利用哪些知觉线索以减少他们在嘈杂环境中的言语识别困难具有重要意义。在多人同时说话的嘈杂环境中老年听者是否能从动态 F0 轮廓线索中获益，目前尚不清楚。

相对于听力正常的年轻人而言，动态 F0 线索对老年听者在噪声下言语感知的影响更为复杂和多变。老年人对 F0 轮廓线索的感知和利用可能会比年轻人更加困难。由于年龄增加所带来的外周听觉系统功能的下降，比如听觉阈限的提高、听觉滤波器带宽的增加、时间分辨率的下降等均可导致老年人的言语加工能力的降低。因此，老年人感知动态 F0 线索的能力可能会因各种阈上听力缺陷而下降，这其中包括时间编码不良(Grose & Mamo, 2010; Hopkins & Moore, 2011)、频率选择性降低(Hopkins & Moore, 2011)，以及频率调制的退化神经表征(Clinard & Cotter, 2015)。这种与年龄相关的听觉时间加工缺陷会影响老年人对语音中时变的基频准确地进行编码，从而影响老年人感知言语中 F0 轮廓线索的能力。

除此之外，老年人高级认知加工功能的下降也会引起他们在嘈杂环境下的言语识别困难(如：Tun et al., 2009; Wingfield & Tun, 2001; Humes et al., 2006)。值得注意的是，纯音听觉阈值并不能完全直接反映老年人的听觉能力(Ruggles et al., 2011)。例如，从感音神经性听力损失和听力正常的老年人中收集的频率辨别数据表明，在听力图相似的老年人中感知音高线索的能力差异很大(Moore & Peters, 1992)。此外，即使拥有临床正常的纯音听觉阈限值，老龄言语感知困难仍可能发生(Füllgrabe et al., 2014)。与年龄相关的噪声中言语理解差异可能是由于与纯音听阈升高无关的其他听觉能力的下降，并涉及中枢听觉或认知处理过程(Füllgrabe et al., 2014)。

认知水平上，老年人在嘈杂环境下处理和理解言语的能力会受到选择性注意、工作记忆、抑制性控制和认知加工速度普遍减慢的影响(如：van Knijff et al., 2018; Schneider et al., 2016)。相对于年轻人，老年人可能需要更多的认知资源，以便在背景噪声存在的情况下解析言语信号，尤其是当输入信号由于听觉时间处理缺陷而进一步消退时，这对自上而下的处理提出了更高的要求。因此，老年人在多人说话的嘈杂环境下处理 F0 轮廓信息会比年轻人更具有挑战性。

虽然老年人外周听觉和中枢听觉处理功能衰退，但有研究证据表明老年人的大脑可能会通过资源重组或认知策略改变以应对感知觉加工能力的下降(Du et al., 2016)。Du 等人(2016)的研究发现老年人在噪声环境下加工和识别言语时会调用更多的认知资源和脑区以作为一种代偿。那么老年人在噪声下感知和利用 F0 线索促进言语识别的能力是否也会存在一定程度的保留？ 老年人因年龄增加带来了听觉加工能力下降，但语言知识实际上是随着年龄的增长而提高的(Wright, 2016)。有研究结果显示在不利的听觉条件下，老年人比年轻人更能从语义语境中受益(Frisina & Frisina, 1997; Sheldon et al., 2008)。当分别从语句、词汇、语音层面给被试提供好的语境条件时，相对于年轻被试，老年被试可以获得更多的目标语句释放量(Sheldon et al, 2008; Singh et al, 2008;Wingfield &Tun, 2007)。此外，已有的证据表明老年听者在安静的环境中对语音理解非常依赖韵律线索(综述见 Wingfield & Tun, 2001)，包括利用动态音高线索来理解语音(Shen & Souza, 2017)。对于在嘈杂环境下难以理解言语的老年人来说，任何对言语感知是冗余的线索也许在不利的环境里有可能可以提供额外的好处(Darwin, 2008)。老年听者也许可以从这些线索中获益，以便更好地识别噪声环境中的语音。

噪声下言语感知能力下降会限制老年人人际交流质量，进而影响到老年人的身心健康。老年人在噪声下言语感知困难背后的机制仍不清楚，治疗方案也很少。动态 F0 轮廓线索对日常交流中的言语识别很重要，听觉加工老龄化是否会影响老年人对动态 F0 轮廓线索的感知和利用，现有的数据集还很有限。F0 轮廓动态变化是否能够帮助老年人从掩蔽中分离出目标言语还有待阐明。嘈杂环境下老年人在汉语言语感知中对 F0 轮廓信息是如何处理的更是知之甚少。值得注意的是，汉语单字的声调具有语义意义，其基频的波动较英语的活跃，这有助于抵抗信息掩蔽。另外，汉语的清辅音较英语的多，会更加受噪音的掩蔽(Kang, 1998)。因此，汉语语境下 F0 轮廓信息在言语感知中是如何处理的需要得到更多关注。本文将基于汉语成句语音，尝试阐明 F0 轮廓线索的操作对言语掩蔽下的汉语言语识别去掩蔽作用受老龄化的影响。

本实验研究将评估老年人与年轻人在言语掩蔽环境下利用汉语成句的 F0 轮廓线索抵抗信息掩蔽识别目标言语的能力，考察对 F0 轮廓线索的操作是否会影响老年人与年轻人在言语掩蔽下的汉语言语识别成绩。通过比较两个年龄组对具有动态自然的 F0 轮廓的语句与具有拉伸的或扁平的 F0 轮廓的相应语句在言语掩蔽下的识别表现，探讨老龄化对 F0 轮廓线索在言语识别去掩蔽作用中的影响。

2 实验方法

2.1 被试

听力正常的 12 名年轻被试(其中 7 名女性和 5 名男性，平均年龄为 24.0 岁，年龄范围在 20~26 岁之间)和 12 名老年被试(其中 6 名女性和 6 名男性，平均年龄为 68.6 岁，年龄范围在 57~75 岁之间)参与了实验[1]。被试的母语均为汉语。所有被试在了解实验目的和程序后都在知情同意书上签字。实验结束后，被试接收适量现金作为实验酬劳。

实验前所有的被试都经过纯音听力测试(如图 1 所示)。年轻被试在 0.125~8 kHz 的频率范围内具有正常(纯音测听阈限不大于 20 dB)且平衡(两耳的阈限差别不大于 15 dB)的纯音测听阈限。老年被试在 0.125~0.5 kHz 的频率范围内具有正常(纯音测听阈限不大于 20 dB)且平衡(两耳的阈限差别不大于 15 dB)的纯音测听阈限，在 1~4 kHz 的频率范围内具有平衡和不大于 45 dB 的纯音测听阈限。不存在传导性听觉损伤。

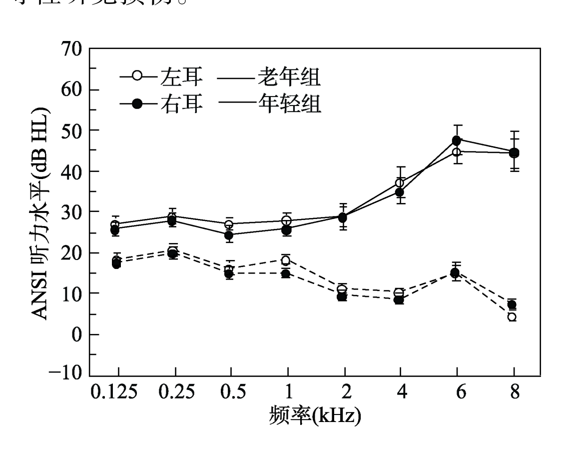


图 1 年轻被试和老年被试的纯音测听阈限

除了以上条件，老年被试还需要通过简易精神状态测试(Mini-mental State Examination, MMSE)和蒙特利尔认知评估(Montreal Cognitive Assessment, MoCA)，要求 MMSE 得分 ≥ 27 且 MoCA 评分 ≥

---

[1]采用 G*Power 3.1.9 软件对研究所需样本量进行先验(priori)分析。以中等效应量 $f = 0.3$，统计检验力 0.95 (α = 0.05) 估计最小样本量为每组 10 个样本(Total sample size:20)，表明本文中当前的被试量(青年组和老年组，每组 12 个)是足够的。

26，以确保参加实验的老年被试都属于认知正常的老年人。

2.2　实验设备

实验在一个安静的语音室中进行。模拟输出信号通过位于水平面正前方的扬声器(Dynaudio Acoustics, BM6 A)呈现给被试。扬声器播出的声音刺激强度在整个实验过程中保持恒定为 60 dBA SPL。实验中，被试坐于语音室正中央的椅子上，面部朝前，保持头部不动，扬声器置于被试正前方位置约为具有平均身高的被试坐位时耳的高度。

2.3　刺激材料

a.**中文“无意义”语句**。实验测试中用到的目标言语刺激为中文无意义句子(Yang et al., 2007)，所有的句子均来自于北京大学言语听觉研究中心自建的语料库，并已被许多研究者 (Chen et al., 2018; Wu, 2019; Wu et al., 2012; Yang et al., 2007)使用。每句中文无意义语句都包含主语(名词)–谓语(动词)–宾语(名词)结构。例如，在一个样例句子“他的磐石可能比赛这个员工”中，下划线的主、谓、宾词为句子的 3 个关键词，每个关键词都是一个双音节词。该无意义语句本身的句子结构对关键词的识别不提供任何上下文信息。

为了确保这些句子都是无意义的，实验对每个句子中所用的两个名词和动词在所选的《人民日报》数据库(1994~2002)中在一个正常句子中同时出现的概率都做了严格限制。只有当无意义语句中的 3 个关键词在所选数据库同一个正常句子中共现的概率为零时，该语句才会被用作实验材料(详见 Yang et al., 2007)。

b.**合成 F0 轮廓语音模型**。为了保证测试句子刺激声学特征的高质量和一致性，本实验使用语音合成技术产生语音材料。目标语音是由机器合成的年轻女性声音朗诵的。语音样本(约 432 句，持续 40 分钟)通过具有特定语音特征的初始声学模型进行训练，以获得不同 F0 轮廓模式的每个目标语音的声学模型。对于每个目标语音，使用生成的目标语音声学模型，可以将无意义的书面体句子转换为具有说话人语音特征(包括基频、特征音高滑动变化、F0 轮廓)的语音信号。最后，将具有不同动态 F0 轮廓特征类型的语音样本加入到初始声学模型中，通过模型训练得到具有不同 F0 轮廓线索(扁平的、自然的、拉伸的 F0 轮廓)的声学语料。

具体地，首先利用层次化音高目标模型和层次化时长模型对目标语句的 F0 轮廓进行建模(Zhang et al., 2010)，获取韵律信息的层次化特征，并对其中音调和语调之间的关联进行量化模拟；接着，分离出音调和句子语调；然后，保留音调，将句调进行参数化操作后，再将其重新叠加、平滑合成为连续的语音。这种操作的目的是在改变句子 F0 轮廓动态变化的同时保持其他韵律线索不变。

使用下列公式对每个目标语句的基频 $f_0(t)$进行参数化操作，为每个句子创建 3 种不同 F0 轮廓动态变化强度的条件：

$$f_0'(t) = [f_0 \times (f_0(t)/f_0)^m] \tag{1}$$

其中，$f_0'(t)$表示调整后的基频轮廓，$f_0(t)$表示目标语音的基频，f_0 表示目标语音基频的平均值，m 是调整参数。自然的 F0 轮廓条件下调整参数设置为 1.0，扁平化的 F0 轮廓条件调整参数设置为 1/4，拉伸的 F0 轮廓条件下设置为 4 (Binns & Culling, 2007; Miller et al., 2010; Grant, 1987; Shen & Souza, 2019; Wu, 2019)。即扁平化的 F0 轮廓条件是把目标句子的 F0 轮廓的偏移量压缩为原来的 1/4，拉伸扩大的 F0 轮廓则将原先 F0 轮廓的偏移量在原来的基础上拉伸放大 4 倍(参见示意图 2)。

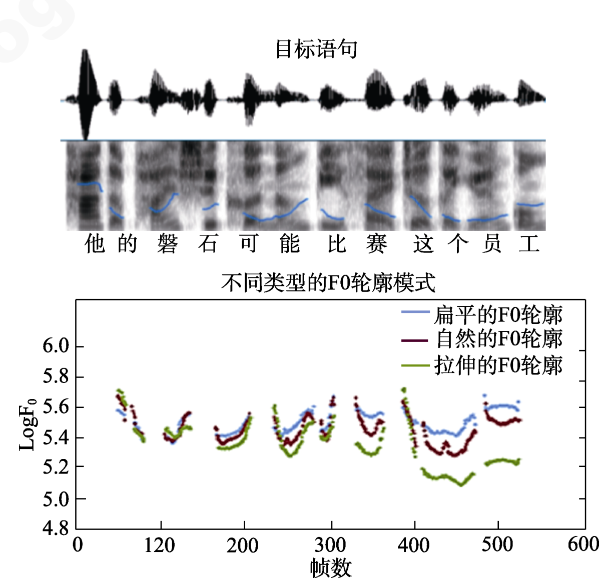


图 2　目标语句的频谱和 F0 轮廓(上)，以及所对应的 3 种不同 F0 轮廓模式(下)的示意图

研究中所选取的对 F0 轮廓处理的参数极值是依据前人的研究结果所选取的(Binns & Culling, 2007; Miller et al., 2010; Grant, 1987)。这些 F0 轮廓处理模式是特别选择的，因为它们在背景噪声中始终显示出对言语感知的影响(Binns & Culling, 2007; Miller et al., 2010; Grant, 1987; Shen & Souza, 2019)，并且这种效应在汉语语言中也得到了证实(Wu, 2019)。

c. **掩蔽刺激**。研究中所用的掩蔽刺激声音是两个与目标声音同性别的说话人声音的等强度叠加形成的言语掩蔽声音(Chen et al., 2018; Yang et al., 2007)。言语掩蔽信号是一个时长为 47 s、对中文无意义语句进行数字合成后得到的连续录音。用于作掩蔽语音的语句由与目标说话人不同的另外两个年轻女性声音说出。所有的言语刺激都被录制成数字化的声音文件，采样率为 22.05 kHz，存储为 16-Bit 的 PCM 波形文件。

2.4 实验设计和程序

实验采用 2 (年龄) × 3 (F0 线索类型) × 4 (信噪比)混合设计。年龄为组间变量，包括年轻和老年被试组。两个被试内变量及其水平分别为：1) F0 线索类型：扁平的 F0 轮廓条件、自然的 F0 轮廓条件、拉伸的 F0 轮廓条件; 2)信噪比：−4, 0, 4, 8 dB SPL。共有 24 种实验条件，每种条件下有 18 个试次。实验中，在每个年龄组内，3 种 F0 线索类型条件在不同的被试间进行完全拉丁方平衡。4 种信噪比在每种 F0 线索类型条件中随机化。

实验中，被试按键开始一个实验试次。按键后，先给被试呈现掩蔽语音，1 s 之后给被试呈现一个目标语句，然后掩蔽语音和目标语句同时结束。被试的任务是在声音结束后立即尽可能地复述所听到的目标句子。实验测试开始前有一个训练阶段，确保被试能够熟悉测试流程。训练阶段用到的目标语句不被用到正式实验中。

3 实验结果

实验过程中对 3 个关键词进行计分，每正确复述一个音节计一分，一个句子满分为 6 分。之后计算每名被试在每种实验条件下(F0 轮廓类型 × 信噪比)的识别率以用于统计分析。

3.1 对正确识别率的数据分析

实验使用 SPSS 统计软件对被试的正确识别百分率进行线性混合效应模型(linear mixed-effects models, LMMs; Baayen et al., 2008)统计分析。在本研究的模型中，包括 F0 轮廓类型(压扁、自然、拉伸)、年龄(年轻、老年)、信噪比水平(−4, 0, 4, 8 dB SPL)的固定效应，同时考虑被试随机效应。

全模型探讨了年龄、F0 轮廓类型、信噪比水平以及这些因素之间的相互作用对听者在言语掩蔽下的汉语言语识别成绩的影响。结果表明，年龄组(F(1, 22) = 194.15, p < 0.001)，F0 线索类型(F(2, 242) = 13.24, p < 0.001)和信噪比(F(3, 242) = 271.91, p < 0.001)的主效应都显著。显著的主效应表明，这三个因素都对汉语言语的可懂度存在贡献。在交互作用方面，F0 线索类型和信噪比两个因素的交互作用显著，F(6, 242) = 2.20, p = 0.043; 信噪比和年龄组两个因素的交互作用显著，F(3, 242) = 10.06, p < 0.001; F0 线索类型和年龄组的交互作用也显著，F(2, 242) = 7.12, p = 0.001，但三因素(F0 线索类型×信噪比×年龄)之间交互作用不显著，F(6, 242) = 1.26, p = 0.272 > 0.1。

在各个年龄组下建立新模型，包括 F0 轮廓类型(压扁、自然、拉伸)、信噪比水平(−4, 0, 4, 8 dB SPL)的主效应和两重交互作用，对实验各因素的效应进行进一步的探测和分析。对年轻被试组的识别成绩统计分析结果表明，对于年轻被试组，F0 线索类型的主效应显著，F(2, 121) = 17.82, p < 0.001; 信噪比的主效应显著，F(3, 121) = 110.11, p < 0.001。F0 线索类型和信噪比两个因素的交互作用显著，F(6, 121) = 2.69, p = 0.017。对老年被试组的识别成绩统计分析结果表明：对老年被试组，信噪比的主效应显著(F(3, 121) = 178.27, p < 0.001)，但是 F0 线索类型的主效应(F(2, 121) = 0.96, p = 0.385)以及 F0 线索类型和信噪比之间的交互作用(F(6, 121) = 0.58, p = 0.744)均不显著。

在各个 F0 轮廓类型条件下建立新模型，包括年龄(年轻、老年)、信噪比水平(−4, 0, 4, 8 dB SPL)的主效应和两重交互作用，对实验各因素的效应做进一步的探测和分析。对于自然 F0 轮廓条件下的识别成绩统计分析结果表明：年龄组的主效应显著，F(1, 22) = 152.07, p < 0.001; 信噪比类型的主效应(F(3, 66) = 73.03, p < 0.001)以及信噪比类型和年龄组之间的交互作用(F(3, 66) = 5.52, p = 0.002)都达到统计显著。对于扁平 F0 轮廓条件下的识别成绩统计分析结果表明：年龄组的主效应显著，F(1, 22) = 117.75, p < 0.001; 信噪比类型的主效应(F(3, 66) = 113.78, p < 0.001)以及信噪比类型和年龄组之间的交互作用(F(3, 66) = 4.58, p = 0.006)均显著。对于拉伸 F0 轮廓条件下的识别成绩统计分析结果表明：年龄组的主效应显著，F(1, 22) = 89.84, p < 0.001; 信噪比类型的主效应显著，F(3, 66) = 81.36, p < 0.001; 但是信噪比类型和年龄组之间的交互作用不显著，F(3, 66) = 2.49, p = 0.065。对两个年龄组被试的识别成绩进一步分析表明，两组被试不管是在自然 F0 轮廓条件下还是在扁平或拉伸 F0 轮廓的条件下对目标语句的识别率都存在显著差异(ps <

0.001)。对年龄组的固定效应参数估计显示, 在每种 F0 轮廓条件下老年被试组的言语识别成绩均显著低于年轻被试组(自然 F0 轮廓条件下: β = 0.26, SE = 0.06, t = 4.70, p < 0.001; 扁平 F0 轮廓条件下: β = 0.17, SE = 0.05, t = 3.20, p = 0.002; 拉伸 F0 轮廓条件下: β = 0.19, SE = 0.06, t = 3.24, p = 0.002)。

在各个信噪比水平条件下建立新模型, 包括年龄(年轻、老年)、F0 轮廓类型(压扁、自然、拉伸)的主效应和两重交互作用, 对实验各因素的效应做进一步的探测和分析。对于信噪比为−4 dB 水平条件下的识别成绩统计分析结果表明: 年龄组的主效应显著, F(1, 22) = 69.48, p < 0.001; F0 类型的主效应显著, F(2, 44) = 7.10, p = 0.002; 以及 F0 类型和年龄组之间的交互作用也显著, F(2, 44) = 4.38, p = 0.018。对于信噪比为 0dB 水平条件下的识别成绩统计分析结果表明: 年龄组的主效应(F(1, 22) = 241.88, p < 0.001), F0 类型的主效应(F(2, 44) = 8.90, p < 0.001), 以及 F0 类型和年龄组之间的交互作用(F(2, 44) = 4.51, p = 0.015)均显著。对于信噪比为 4 dB 水平条件下的识别成绩统计分析结果表明: 年龄组的主效应显著, F(1, 22) = 53.89, p < 0.001; F0 类型的主效应显著, F(2, 44) = 4.85, p = 0.013; 但 F0 类型和年龄组之间的交互作用不显著, F(2, 44) = 0.83, p = 0.441。对于信噪比为 8dB 水平条件下的识别成绩统计分析结果表明: 年龄组的主效应显著, F(1, 22) = 34.11, p < 0.001; 但 F0 类型的主效应(F(2, 44) = 0.80, p = 0.458)以及 F0 类型和年龄组之间的交互作用(F(2, 44) = 1.30, p = 0.283)都不显著。

3.2　对识别阈限值的数据分析

为了更清楚揭示年龄组和 F0 线索类型对被试识别成绩的影响, 我们用公式(2)中的心理测量参数方程对每名被试的识别率数据进行了曲线拟合, 得到每名被试在每种条件下的识别阈限值用于数据分析:

$$y = 1 / [1 + e^{-\sigma(x-\mu)}] \quad (2)$$

式中, y 是关键词的正确识别率, x 是与 y 对应的信噪比, μ 是在曲线上对应于 50%识别率处的信噪比, 而 σ 则是心理测量函数的斜率。参数 μ 和 σ 通过 Levenberg-Marquardt 方法拟合得到。

通过参数方程拟合可以得到每名被试在 3 种不同 F0 线索条件下的识别阈限值(达到 50%正确识别率时所对应的信噪比值)。图 3 给出了言语掩蔽下老年人与年轻人在 3 种不同 F0 线索条件下的识别率及相应的参数拟合曲线。

实验使用 SPSS 统计软件, 采用线性混合效应

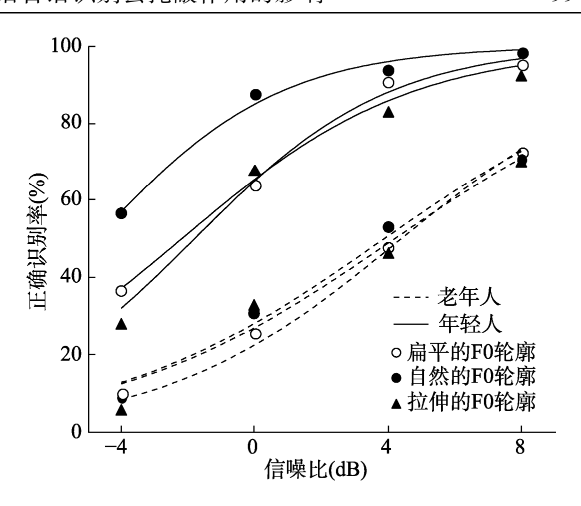


图 3　不同 F0 轮廓条件下识别率随信噪比变化的函数以及对应的参数方程拟合曲线

注: 图中用实线表示年轻人的识别成绩, 平滑的曲线是对数据进行参数拟合的结果。虚线部分显示的是老年人的识别成绩进行参数拟合的结果。图中 3 种标记分别描述 3 种不同 F0 轮廓条件下的识别率: 1)扁平的 F0 轮廓条件(空心圆); 2)自然的 F0 轮廓条件(实心圆); 3)拉伸的 F0 轮廓条件(实心三角形)。

模型对被试的识别阈限值进行统计分析。在本研究的模型中, 包括 F0 轮廓类型条件(压扁、自然、拉伸)、年龄组(年轻、老年)的固定效应, 同时考虑被试随机效应。

实验分析结果显示, F0 轮廓类型(F(2, 44) = 7.15, p = 0.002)和年龄组组间差异(F(1, 22) = 73.21, p < 0.001)对被试的识别成绩均有显著影响, F0 线索类型和年龄组的交互作用显著, F(2, 44) = 4.93, p = 0.012。固定效应参数估计进一步显示, 老年被试组的言语识别成绩明显差于年轻被试组(β = −7.73E0, SE = 0.90, t = −8.56, p < 0.001)。

来自年轻被试组识别阈限值的数据分析显示, 在对 F0 线索操作的影响方面, 不同 F0 轮廓类型对被试识别成绩的影响显著, F(2, 22) = 6.87, p = 0.005。在模型中将自然 F0 轮廓条件作为基线条件, 对 F0 轮廓被调节的条件和自然 F0 轮廓条件之间的识别成绩进行比较。结果显示, 与自然动态 F0 轮廓线索条件相比, 压扁 F0 轮廓条件下的言语识别表现明显较差(β = 2.92E0, SE = 0.88, t = 3.33, p = 0.003); 拉伸 F0 轮廓条件下的言语识别成绩也比自然动态 F0 轮廓条件下的显著较低(β = 2.70E0, SE = 0.88, t = 3.08, p = 0.005)。但压扁 F0 轮廓与拉伸 F0 条件下的言语识别成绩之间差异不显著(β = 0.21, SE = 0.88, t = 0.24, p = 0.811)。

对老年被试组识别阈限值的数据分析表明, 不同 F0 轮廓线索类型对老年被试的言语识别成绩的

影响不显著, $F(2, 22) = 0.50$, $p = 0.614$。压扁 F0 轮廓条件($\beta = 0.34$, $SE = 0.34$, $t = 0.99$, $p = 0.33$)或拉伸 F0 轮廓条件($\beta = 0.16$, $SE = 0.34$, $t = 0.48$, $p = 0.637$)下的言语识别成绩与自然动态 F0 轮廓线索条件下的言语识别成绩之间差异均不显著。

为了进一步探究数据支持两个主效应以及主效应间交互作用显著的可靠性, 使用统计软件 JASP (https://jasp-stats.org/, JASP Team 2019) (Wagenmakers et al., 2018)对数据进行贝叶斯分析检验, 分析产生的贝叶斯因子(BF_{10})描述了数据对假设的支持程度, 能够量化备择假设 H_1 和零假设 H_0 的相对证据。

实验使用 JASP 的默认分布, 对两个年龄被试组的识别阈限值数据进行贝叶斯重复测量方差分析, 结果显示年龄组主效应的 BF_{10} 值是 635658.21, 根据贝叶斯因子的决策标准(Jeffreys, 1961), 该结果表明有极强的证据支持 H_1; F0 轮廓类型主效应的 BF_{10} 值是 8.66, 有中等强度的证据支持 H_1; 年龄组和 F0 轮廓类型的主效应之和的 BF_{10} 值是 5.59e+6, 表明与单独的 F0 轮廓主效应相比, 当加入年龄的效应时, 数据对 H_1 的支持提高了大约 6.46e+05 倍; 在此基础上, 当把年龄组和 F0 轮廓类型的交互作用也增加到模型中时, BF_{10} 值是 2.25e+7, 表明有极强的证据支持 H_1。

对年轻被试组的数据进行贝叶斯重复测量方差分析, 结果显示贝叶斯因子 $BF_{10} = 12.32$; 说明有较强的证据支持 H_1, 即年轻人 F0 轮廓类型的主效应显著。对老年被试组的数据进行贝叶斯重复测量方差分析, 结果显示贝叶斯因子 $BF_{10} = 0.27$; 说明存在中等强度的证据支持 H_0, 即老年人的 F0 轮廓线索类型的主效应不显著。

进一步的简单效应检验使用 JASP 默认分布对老年被试在不同 F0 轮廓类型条件下的数据两两进行贝叶斯配对样本 t 检验, 分析得出的贝叶斯因子数值(压扁的 F0 轮廓条件和自然 F0 轮廓条件, BF_{10} 为 0.42; 自然 F0 轮廓条件和拉伸 F0 轮廓条件, BF_{10} 为 0.34; 压扁 F0 轮廓条件和拉伸 F0 轮廓条件, BF_{10} 为 0.32), 根据 Jeffreys (1961)提出的分类标准, 表明当前的数据有较弱的证据支持了 H_0, 也说明观察到的老年被试的数据相对于备择假设 H_1 更加支持没有效应的假设 H_0, 即老年被试的识别成绩在不同 F0 轮廓类型的条件之间没有显著差异。

4 讨论

本研究考察了 F0 轮廓线索的操作对老年人与年轻人在言语掩蔽下的汉语言语识别成绩的影响。研究结果表明老龄化会影响老年听者对动态 F0 轮廓线索的感知和利用。实验结果发现所有年龄组在不同 F0 线索条件下的言语识别成绩都会随着信噪比的增加而提高, 研究结果也证实了动态 F0 轮廓线索有助于年轻人在言语掩蔽下有效的语音识别, 然而对于听力正常的老年人来说, 这种动态 F0 轮廓线索的获益似乎是有限的。研究结果显示, 在多人说话的言语掩蔽下, F0 轮廓自然的目标比扁平或拉伸的 F0 轮廓的目标更能帮助年轻被试抵抗掩蔽; 而相比年轻人, 老年人在嘈杂环境中很可能遵循 F0 轮廓动态变化出现了困难, 使得老年人难以利用 F0 轮廓线索提高对背景掩蔽声中的目标言语识别, 从而导致 F0 轮廓动态变化促进老年人从掩蔽语音流中分离出目标言语的作用变得有限。

实验研究评估了老年被试组与年轻被试组在言语掩蔽下聆听具有自然 F0 轮廓的汉语语句与拉伸的或扁平的 F0 轮廓的相应语句的言语识别成绩。数据分析结果发现 F0 轮廓类型、信噪比水平和年龄因素都会对听者言语识别成绩产生影响。研究结果显示信噪比对年轻被试和老年被试都存在显著的影响($ps < 0.001$)。不管是老年被试组还是年轻被试组, 在言语掩蔽下关键词的识别率都随着信噪比的提高而升高。具体地, 在本实验的所有条件中, 关键词的识别率都随着信噪比从−4 dB 增加到 8 dB 而呈单调增加的趋势, 没有表现出任何平台效应。言语掩蔽的干扰声音下, 这种目标识别率的单调递增趋势是与前人所报告的结果一致(Brungart et al., 2001; Li et al., 2004; Wu, 2019)。

研究结果还显示, 对于年轻被试组, 不管是在自然 F0 轮廓条件下还是在扁平或拉伸 F0 轮廓的条件下, 年轻被试对目标语句的识别率都随信噪比的降低而降低。数据分析结果表明在每种 F0 轮廓条件下, 信噪比对目标识别率的作用都显著($ps < 0.001$)。对于老年被试组, 实验结果也显示不管在自然 F0 轮廓条件下还是在扁平或拉伸 F0 轮廓的条件下, 老年被试对目标语句的识别率都随信噪比的变化而变化。数据分析表明信噪比对识别率的作用不管是在自然 F0 轮廓条件($p < 0.001$)下还是在被调节的 F0 轮廓线索的条件($ps < 0.001$)下都显著。这个结果与年轻被试的结果一致。然而, 与年轻被试的结果不同的是, 对老年被试而言, 自然动态 F0 轮廓条件产生的去掩蔽作用与扁平的 F0 轮廓条件或拉伸的 F0 轮廓条件产生的去掩蔽作用相当($ps >$

0.05)。而年轻被试在自然动态 F0 轮廓条件产生的去掩蔽作用与扁平的 F0 轮廓条件(p < 0.001)或拉伸的 F0 轮廓条件(p < 0.001)产生的去掩蔽作用存在显著差异。

具体实验测量结果显示, 言语掩蔽下给年轻被试呈现中文无意义语句时, 具有自然动态 F0 轮廓的语句比具有扁平的 F0 轮廓或拉伸 F0 轮廓的语句都表现出较好的言语识别成绩。说明具有自然 F0 轮廓的条件比具有拉伸或扁平 F0 轮廓的条件更好地被感知。这一结论发现证明了自然 F0 轮廓线索对年轻听者在噪声下言语识别中的重要作用。研究结果也表明了对 F0 轮廓线索的操作会对听者在言语掩蔽下的汉语言语识别表现产生有害影响。本研究的结果与英语语音感知研究的文献(Binns & Culling, 2007; Miller et al., 2010; Shen & Souza, 2019)中的结论一致, 表明与自然的 F0 轮廓模式相比, 被调节操作的动态 F0 线索对噪声中的言语感知有负面影响。同时, 本研究结果也支持了 Patel 等人(2010)和 Wu (2019)中关于 F0 轮廓的自然动态变化对背景噪声中的汉语言语识别具有重要贡献的论述。

对老年被试的数据分析结果显示, 老年被试在言语掩蔽下聆听具有自然 F0 轮廓的汉语语句与拉伸的或扁平的 F0 轮廓的相应句子呈现出同样低的言语识别成绩。研究结果表明, 扁平或拉伸 F0 轮廓的条件没有能帮助老年人在言语掩蔽环境下抵抗信息掩蔽, 自然动态变化的 F0 轮廓的条件也没有比对 F0 轮廓操作后的条件更能抵抗信息掩蔽, 说明在多人说话的嘈杂环境下老年听者可能难以利用动态 F0 轮廓线索以促进对掩蔽下的目标言语识别。

在多人同时说话的嘈杂环境下, 听者会在知觉层次上利用一些线索来促进他们在目标语音和掩蔽语音之间的知觉分离, 提高对目标言语的有效识别(Schneider et al., 2016)。嘈杂环境下年轻听者可以利用语音基频的差异来区分目标语音和掩蔽语音(Darwin et al., 2003)。当目标语音和掩蔽语音同时存在时, 目标语与掩蔽语之间的 F0 差异能够加强听者对目标言语的跟踪, 进而提高言语识别率(Brungart et al., 2001)。然而老龄化会对处理竞争声音之间的基频差异的能力造成有害影响。有研究证据显示老年人在处理安静环境中声音的基频信息时有更大的困难(Lee, 2013)。老年被试的 F0 辨别阈值明显大于年轻被试(Souza et al., 2011; Vongpaisal & Pichora-Fuller, 2007); 老年人也表现出使用 F0 差异来感知竞争声的难度增加(如: Rossi-Katz & Arehart, 2009; Vongpaisal & Pichora-Fuller, 2007)。由于老年人在处理相互竞争的说话者之间的 F0 差异方面增加了难度, 因此 F0 处理的缺陷可能会导致老年人在嘈杂环境下从 F0 线索中获益的能力存在年龄相关性下降。

言语的时间结构分成为周期性线索、包络线索和精细结构信息三个不同的水平, 这几种时域线索相互补充, 传递了语音的完整信息(Rosen, 1992)。给听力正常的听者播放言语语音时, F0 感知的主导线索是分解的低频谐波(Arehart, 1994; Oxenham, 2008), 即精细结构; 此外 F0 感知也可以从高频未分解谐波传递的高速包络波动中得出, 即周期性(Oxenham, 2008)。许多证据表明老年人比年轻人在有言语掩蔽下的语音识别任务以及涉及周期性线索的心理物理任务方面表现更差(如: Vongpaisal & Pichora-Fuller, 2007; Arehar et al., 2011)。这些研究中普遍存在的假设是, 周期性线索有助于促进语音流分离, 而老年人比年轻人更难意识到这一好处(Bologna et al., 2019)。已有的研究证明了在多人说话的言语识别任务中, 为 F0 提供周期性线索可以提高句子识别能力(Stone et al., 2008)。然而老龄化会对周期性和精细结构知觉产生不利影响。例如, 在脑干周期性线索的神经表征中观察到了年龄相关的下降(Clinard & Tremblay, 2013; Snyder & Alain, 2005)。老年人难以使用周期性线索来获取与说话者语音基本频率相关的信息(Souza et al., 2011)。因此, 与年龄相关的周期性编码下降会对老年人受益于 F0 轮廓线索的能力产生不利影响。

此外, 由于年龄相关的时域包络(如: Souza & Boike 2006; Souza et al. 2011)和精细结构(e.g. Vongpaisal & Pichora-Fuller 2007; Hopkins & Moore 2011; Smith et al., 2002)线索处理方面的缺陷, 老年人可能不能像年轻人一样使用和整合不同类型的线索进行言语识别。基频(F0)轮廓线索对老年人在目标言语识别中的贡献模式可能也不同于年轻人。包络成分和精细结构成分信息对听觉感知具有不同的贡献。精细结构主要对语音的音调识别和声源定向起重要作用, 而包络主要对言语的内容识别起重要作用(李量 等, 2017; Smith et al., 2012)。基频 F0 是一种精细结构的线索(Brown & Bacon, 2010), 有利于在噪声中的言语识别。然而老龄化会降低对精细结构的感知敏感度, 从而影响对精细结构信息的编码能力(Pichora-Fuller et al., 2007)。这种对声音精细结构信息处理精确性的缺失会使得老年人在

需要跟踪基频变化的情况下处于不利地位。因此，与年轻人相比，老年人在言语掩蔽下遵循 F0 轮廓动态变化更为困难。

本文的研究结果提供证据表明了老龄化会影响听者对动态 F0 轮廓线索的感知和利用。与自然动态 F0 轮廓条件相比，本项研究中当前的 F0 轮廓调节策略并不有利于老年人在言语掩蔽下感知目标语音。这种不利可能源于多种因素，包括背景掩蔽类型和言语刺激选择的影响，这些都有待于今后的研究阐明。例如，本文研究中使用的是汉语语句，言语信号在频谱结构和时间包络上快速不断地变化，这些复杂的模式很可能会干扰老年人对动态 F0 轮廓线索的感知和利用。因此，在不利条件下，错过这些线索的受益可能会增加老年人的言语识别难度。语音信号是调幅调频信号，频谱线索由信号的谐波结构、基频变化和共振峰共同构成。共振峰的一个重要特性在于它们在背景噪声中可以得到比较好的保持(Assmann & Summerfield, 2004)。共振峰频率的动态模式会伴随着信号的频谱和时间包络线而变化，这可能会模糊听者感知 F0 轮廓动态变化所需的线索(Souza et al., 2011)。所以，老年听者感知并追踪动态 F0 轮廓变化出现困难可能是因为对共振峰变化的干扰更敏感，从而对他们在言语识别中从动态 F0 线索中获益造成不利影响，值得未来研究的进一步关注。此外，本研究所使用的言语刺激是无意义的中文句子。应该探索的一种可能性是，为了提高背景掩蔽声下的言语可懂度，动态 F0 线索是否可以与其他线索协同工作。未来的研究可以揭示这种可能性，研究聚焦于探讨那些具有语言学意义的 F0 线索的增强策略的方法是否能够改善老年听者在背景掩蔽声下的言语感知。

值得注意的是，在前人的研究中，老年人年龄相关性听力下降与声音辨别能力(Helfer & Freyman, 2008)或声音精细结构保持的初始听觉“记忆”能力(Huang et al., 2008)没有显著相关性；老年人在利用知觉空间分离线索从信息掩蔽中释放目标语音的能力并不比年轻人差(Helfer & Freyman, 2008; Li et al., 2004)。尽管如此，不应完全排除年龄相关性听力损失在损害利用 F0 轮廓线索以促进噪声下言语感知方面的作用。在未来的工作中，研究阈上听觉能力测量和动态 F0 线索获益之间的关系具有重要意义。从临床角度来看，这项工作还可以扩展到包括一组患有听力损失的年轻人或老年人，以研究听力损失对噪声下言语识别的动态 F0 线索获益程度的影响。

5 结论

老年人在安静环境下的言语知觉和理解能力正常，但在日常交流的嘈杂环境中理解言语时总是比年轻人遇到更多的困难。基频(F0)轮廓有助于日常交流中的言语识别，可以作为从背景噪声中分离语音的感知线索。本文考察了老年人与年轻人在言语掩蔽下聆听具有自然 F0 轮廓的汉语语句与拉伸的或扁平的 F0 轮廓的相应语句的言语识别能力。研究结果表明了句子 F0 轮廓信息对听者在噪声下汉语言语感知具有重要作用。实验结果显示年轻人在言语掩蔽环境下能够利用动态 F0 轮廓线索抵抗信息掩蔽识别目标言语。而相比年轻人，老年人在多说话人环境下的汉语言语识别中似乎遵循 F0 轮廓动态变化出现困难。这种感知和追踪 F0 轮廓动态变化能力的年龄相关性下降，使得言语掩蔽下动态 F0 轮廓线索对老年人的汉语言语识别去掩蔽作用变得有限。研究结果也体现了汉语成句 F0 轮廓线索对老年人和年轻人在言语掩蔽下目标识别中贡献模式的差异性。本项研究尝试通过揭示老年人利用 F0 轮廓线索促进噪声下汉语言语感知能力的老化特点，为老年听者在嘈杂环境下往往难以理解言语提供一种解释。

参 考 文 献

Arehart, K. H. (1994). Effects of harmonic content on complex-tone fundamental-frequency discrimination in hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *95*(6), 3574–85.

Arehart, K. H., Souza, P. E., Muralimanohar, R. K., & Miller, C. W. (2011). Effects of age on concurrent vowel perception in acoustic and simulated electroacoustic hearing. *Journal of Speech Language and Hearing Research, 54*(1), 190–210.

Assmann, P. F., Summerfield, Q. (2004). The perception of speech under adverse conditions. In S. Greenberg, W. A. Ainsworth, A. N. Popper, & R. R. Fay (Eds.), *Springer Handbook of Auditory Research: Vol. 14: Speech processing in the auditory system* (pp. 231–308). New York: Springer.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412.

Binns, C., & Culling, J. F. (2007). The role of fundamental frequency contours in the perception of speech against interfering speech. *The Journal of the Acoustical Society of America*, *122*(3), 1765–1776.

Bologna, W. J., Vaden, K. I., Ahlstrom, J. B., & Dubno, J. R. (2019). Age effects on the contributions of envelope and periodicity cues to recognition of interrupted speech in quiet and with a competing talker. *The Journal of the Acoustical Society of America*, *145*(3), EL173–EL178.

Brown, C. A., & Bacon, S. P. (2010). Fundamental frequency and speech intelligibility in background noise. *Hearing*

Research, *266*(1-2), 52–59.

Brungart, D. S., Simpson, B. D., Ericson, M. A., & Scott, K. R. (2001). Informational and energetic masking effects in the perception of multiple simultaneous talkers. *The Journal of the Acoustical Society of America, 110*(5), 2527–2538.

Chen, J., Yang, H. Y., Wu, X. H., & Moore, B. C. J. (2018). The effect of F0 contour on the intelligibility of speech in the presence of interfering sounds for Mandarin Chinese. *The Journal of the Acoustical Society of America, 143*(2), 864–877.

Clinard, C. G., & Cotter, C. M. (2015). Neural representation of dynamic frequency is degraded in older adults. *Hearing Research, 323*, 91–98.

Clinard, C. G., & Tremblay, K. L. (2013). Aging degrades the neural encoding of simple and complex sounds in the human brainstem. *Journal of the American Academy of Audiology*, *24*(7), 590–644.

Cutler, A., Dahan, D., & van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech*, *40*, 141–201.

Dannenbring, G. L. (1976). Perceived auditory continuity with alternately rising and falling frequency transitions. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, *30*(2), 99–114.

Darwin, C. J. (2008). Listening to speech in the presence of other sounds. *Philosophical Transactions of the Royal Society B*, *363*(1493)*,* 1011–1021.

Darwin, C. J., Brungart, D. S., & Simpson, B. D. (2003). Effects of fundamental frequency and vocal-tract length changes on attention to one of two simultaneous talkers. *The Journal of the Acoustical Society of America, 114*(5), 2913–2922.

Du, Y., Buchsbaum, B. R., Grady, C. L., & Alain, C. (2016). Increased activity in frontal motor cortex compensates impaired speech perception in older adults. *Nature Communications, 7,* 12241.

Frisina, D. R., & Frisina, R. D. (1997). Speech recognition in noise and presbycusis: Relations to possible neural mechanisms. *Hearing Research*, *106*(1-2), 95–104.

Füllgrabe, C., Moore, B. C., & Stone, M. A. (2014). Age-group differences in speech identification despite matched audiometrically normal hearing: contributions from auditory temporal processing and cognition. *Frontiers in Aging Neuroscience*, *6*, 347.

Grant, K. W. (1987). Identification of intonation contours by normally hearing and profoundly hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *82*(4), 1172–1178.

Greenberg, S., Ainsworth, W., Popper, A., & Fay, R. (2004). *Speech processing in the auditory system*: Springer Press.

Grose, J. H., & Mamo, S. K. (2010). Processing of temporal fine structure as a function of age. *Ear and Hearing*, *31*(6), 755–760.

Helfer, K. S., & Freyman, R. L. (2008). Aging and speech-on-speech masking. *Ear and Hearing*, *29*(1), 87–98.

Hillenbrand, J. M. (2003). Some effects of intonation contour on sentence intelligibility. *The Journal of the Acoustical Society of America*, *114*(4), 2338.

Hopkins, K., & Moore, B. (2011). The effects of age and cochlear hearing loss on temporal fine structure sensitivity, frequency selectivity, and speech reception in noise. *The Journal of the Acoustical Society of America*, *130*(1), 334-349.

Huang, Y., Huang, Q., Chen, X., Qu, T. S., Wu, X. H., & Li, L. (2008). Perceptual integration between target speech and target-speech reflection reduces masking for target-speech recognition in younger adults and older adults. *Hearing Research*, *244*(1-2), 51–65.

Humes, L. E., Lee, J. H., & Coughlin, M. P. (2006). Auditory measures of selective and divided attention in young and older adults using single-talker competition. *Journal of the Acoustical Society of America*, *120*(5), 2926–2937.

Jeffreys, H. (Ed.). (1961). *Theory of probability* (3rd ed.). Oxford, UK: Oxford University Press.

Kang, J. (1998). Comparison of speech intelligibility between english and chinese. *The Journal of the Acoustical Society of America*, *103*(2), 1213–1216.

Lee, J. H. (2013). Age-related deficits in the processing of fundamental frequency differences for the intelligibility of competing voices. *Korean Journal of Audiology*, *17*(1), 1–8.

Li, L., Daneman, M., Qi, J. G., & Schneider, B. A. (2004). Does the information content of an irrelevant source differentially affect spoken word recognition in younger and older adults? *Journal of Experimental Psychology: Human Perception and Performance, 30*(6), 1077–1091.

Li, L., Zheng, Y. J., Wu, C, Li, J. H., Zhang, C. X., & Lu, L. X. (2017). The brain network mechanisms underlying perceptual unmasking cue-induced improvement of speech recognition under cocktail-party listening conditions. *Advances in Psychological Science*, *25*(12), 2099–2110.

[李量, 郑英君, 吴超, 黎绢花, 张畅芯, 陆灵犀. (2017). 在鸡尾酒会场景下利用去掩蔽知觉线索提高言语识别的脑网络机制. *心理科学进展*, *25*(12), 2099–2110.]

Miller, S. E., Schlauch, R. S., & Watson, P. J. (2010). The effects of fundamental frequency contour manipulations on speech intelligibility in background noise. *The Journal of the Acoustical Society of America, 128*(1), 435–443.

Moore, B., & Peters, R. W. (1992). Pitch discrimination and phase sensitivity in young and elderly subjects and its relationship to frequency selectivity. *The Journal of the Acoustical Society of America*, *91*(5), 2881.

Oxenham, A. J. (2008). Pitch perception and auditory stream segregation: implications for hearing loss and cochlear implants. *Trends in Amplification*, *12*(4), 316–331.

Patel, A. D., Xu, Y., & Wang, B. (2010, May). *The role of F0 variation in the intelligibility of Mandarin sentences.* Paper presented at the International Conference on Speech Prosody, 2010, Chicago, USA.

Pichora-Fuller, M. K., Schneider, B. A., Macdonald, E., Pass, H. E., & Brown, S. (2007). Temporal jitter disrupts speech intelligibility: A simulation of auditory aging. *Hearing Research, 223*(1-2), 114–121.

Rosen, S. (1992). Temporal information in speech: acoustic, auditory and linguistic aspects. *Philosophical Transactions of The Royal Society B Biological Sciences*, *336*(1278), 367–373.

Rossi-Katz, J., & Arehart, K. H. (2009). Message and talker identification in older adults: Effects of task, distinctiveness of the talkers' voices, and meaningfulness of the competing message. *Journal of Speech Language and Hearing Research*, *52*(2), 435–453.

Ruggles, D., Bharadwaj, H., & Shinn-Cunningham, B. G. (2011). Normal hearing is not enough to guarantee robust encoding of suprathreshold features important in everyday communication. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(37), 15516–15521.

Schneider, B. A., Avivi-Reich, M., & Daneman, M. (2016). How spoken language comprehension is achieved by older listeners in difficult listening situations. *Experimental Aging Research, 42*(1), 40–63.

Schneider, B. A., Li, L., & Daneman, M. (2007). How competing speech interferes with speech comprehension in everyday listening situations. *Journal of the American Academy of Audiology, 18*(7), 559–572.
Sheldon, S., Pichora-Fuller, M. K., & Schneider, B. A. (2008). Priming and sentence context support listening to noise-vocoded speech by younger and older adults. *The Journal of the Acoustical Society of America*, *123*(1), 489–499.
Shen, J., & Souza, P. E. (2017). Do older listeners with hearing loss benefit from dynamic pitch for speech recognition in noise?. *American Journal of Audiology*, *26*(3S), 462–466.
Shen, J., & Souza, P. E. (2019). The ability to glimpse dynamic pitch in noise by younger and older listeners. *The Journal of the Acoustical Society of America, 146*(3), El232–El237.
Singh, G., Pichora-Fuller, M. K., & Schneider, B. A. (2008). The effect of age on auditory spatial attention in conditions of real and simulated spatial separation. *The Journal of the Acoustical Society of America*, *124*(2), 1294–1305.
Smith, Z. M., Delgutte, B., & Oxenham, A. J. (2002). Chimaeric sounds reveal dichotomies in auditory perception. *Nature*, *416*(6876), 87–90.
Snyder, J. S., & Alain, C. (2005). Age-related changes in neural activity associated with concurrent vowel segregation. *Cognitive Brain Research, 24*(3), 492–499.
Souza, P., Arehart, K., Miller, C. W., & Muralimanohar, R. K. (2011). Effects of age on F0 discrimination and intonation perception in simulated electric and electroacoustic hearing. *Ear and Hearing*, *32*(1), 75–83.
Souza, P. E., & Boike, K. T. (2006). Combining temporal-envelope cues across channels: Effects of age and hearing loss. *Journal of Speech, Language, and Hearing Research*, *49*(1), 138–149.
Stone, M. A., Füllgrabe, C., & Moore, B. (2008). Benefit of high-rate envelope cues in vocoder processing: Effect of number of channels and spectral region. *The Journal of the Acoustical Society of America*, *124*(4), 2272–82.
Tun, P. A., McCoy, S., & Wingfield, A. (2009). Aging, hearing acuity, and the attentional costs of effortful listening. *Psychology and Aging*, *24*(3), 761–766.
van Knijff, E. C., Coene, M., & Govaerts, P. J. (2018). Speech understanding in noise in elderly adults: The effect of inhibitory control and syntactic complexity. *International Journal of Language and Communication Disorders*, *53*(3), 628–642.
Vongpaisal, T., & Pichora-Fuller, M. K. (2007). Effect of age on F0 difference limen and concurrent vowel identification. *Journal of Speech, Language, and Hearing Research*, *50*(5), 1139–1156.
Wagenmakers, E.-J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., ... Morey, R. D. (2018). Bayesian inference for psychology. Part II: Example applications with JASP. *Psychonomic Bulletin & Review*, *25*(1), 58–76.
Wang, J., Hua, S., Zhang, L., Liu, Z., & Zhang, Y. (2013). The roles of fundamental frequency contours and sentence context in mandarin chinese speech intelligibility. *The Journal of the Acoustical Society of America*, *134*(1), EL91–7.
Wingfield, A., & Tun, P. A. (2001). Spoken language comprehension in older adults: Interactions between sensory and cognitive change in normal aging. *Seminars in Hearing*, *22*(3), 287–302.
Wingfield, A., & Tun, P. A. (2007). Cognitive supports and cognitive constraints on comprehension of spoken language. *Journal of the American Academy of Audiology*, *18*(7), 548–558.
Wright, H. H. (Ed.). (2016). *Cognition, language and aging*. Amsterdam: John Benjamins Publishing Company.
Wu, M. H. (2019). Effect of F0 contour on perception of Mandarin Chinese speech against masking. *Plos One, 14*(1), e0209976.
Wu, M. H., Li, H. H., Hong, Z. L., Xian, X. C., Li, J. Y., Wu, X. H., & Li, L. (2012). Effects of aging on the ability to benefit from prior knowledge of message content in masked speech recognition. *Speech Communication, 54*(4), 529–542.
Xu, L., & Zhou, N. (2011). Tonal languages and cochlear implants. In: F. G. Zeng, A. N. Popper, & R. R. Fay (Eds.), *Auditory prostheses: New horizons* (pp. 341–364). Springer; New York.
Yang, Z. G., Chen, J., Huang, Q., Wu, X. H., Wu, Y. H., Schneider, B. A., & Li, L. (2007). The effect of voice cuing on releasing Chinese speech from informational masking. *Speech Communication, 49*(12), 892–904.
Zhang, Z. P., Wang, X., Yu, Y. S., & Wu, X. H. (2010, November). Hierarchical pitch target model for Mandarin speech. *IEEE 2010 7th International Symposium on Chinese Spoken Language Processing* (pp.378–382).

The influence of aging on the unmasking effect of F0 contour cue in Chinese speech recognition

WU Meihong

(*School of Informatics, Xiamen University; Institute of Psychology, Xiamen University, Xiamen* 361005, *China*)

Abstract

Older adults encounter difficulty in recognizing speech in environments where multiple people are talking. Fundamental frequency (F0) contour is very important for speech recognition in daily communication and can serve as a perceptual cue to segregate speech from background noise. The effect of dynamic F0 contour cues on the speech recognition of younger adults in noisy environments has been widely studied, but the influence on older adults' speech recognition, especially in tonal languages like Chinese, is still unclear.

The present study explores whether older adults can benefit from dynamic F0 contour cues for Chinese

speech recognition under the masking of speech with 12 elderly participants (6 female and 6 male, mean age 68.6 years) and 12 young participants (7 male and 5 female, aged 18~25 years old) possessing normal peripheral hearing. The speech recognition threshold for natural F0 contour sentences and the corresponding sentences with F0 contour manipulations (flattened vs. exaggerated) under two-speaker anomalous speech masking for younger adults and older adults were measured, as well as their ability to follow the target speech under speech masking.

The results showed that natural dynamic F0 contour cues are important to Chinese speech recognition when multiple people are speaking. Speech recognition in all age groups improved with higher signal-to-noise ratios, but while natural F0 patterns facilitated performance of younger adults under speech masking, it had no significant effect on older adults. Under masking with two speakers, the natural dynamic F0 contour cues assist young adults to resist information masking more than a flattened or exaggerated F0 contour, but for the older adults, the speech intelligibility of the target sentences with a natural F0 contour was as poor as that of target sentences with flattened or exaggerated F0 contour.

The availability of F0 contour cues seriously affects the benefits of dynamic F0 contour cues for older adults' speech recognition against speech masking. There also appears to be an age-related reduction of the benefit from dynamic F0 contour cues in masked speech recognition, so the F0 contour of Chinese sentences may contribute more to speech recognition under speech masking for younger adults than for older adults.

Keywords Chinese speech recognition, fundamental frequency contour, age-related deficits, speech masking, unmasking